Aalto University
School of Business

FORECASTING STOCK PRICE CRASHES IN THE BIOTECHNOLOGY SECTOR

Evidence from the U.S. in 2002-2016

Master's Thesis
Sanni Virtanen
Aalto University School of Business
Accounting
Fall 2017

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Sanni Virtanen	
Title of thesis Forecasting stock price crashes in the biotechnology sector – Evidence from the U.S. in 2002-2016	
Degree Master of Science in Economics and Business Administration	
Degree programme Accounting	
Thesis advisor Henry Jarva	
Year of approval 2017	**Number of pages** 54 **Language** English

Abstract

This study examines determinants for stock price crashes in the U.S. biotechnology sector. Further, the study investigates three crash metrics' consistency in firm-specific crash classification.

The study is quantitative in nature and applies both logistic and multiple regression to explore the ability of a set of chosen explanatory variables to predict crashes. Two existing crash risk metrics are used, and one additional metric is created as a part of the study. The sample consist of 1303 observations from the U.S. biotechnology/pharmaceutical industry in 2002-2016.

The results of this study suggest that volatility of a stock return is positively associated with future crash risk, and return on assets as well as firm size in terms of total assets and market capitalization are negatively associated with future stock price crash risk. In relation to the consistency of crash classifications, the results show that 20-30% of the classifications are inconsistent within the results of the three crash metrics used in this study.

Keywords stock price crash risk, biopharmaceutical industry, biotechnology, pharmaceutical industry, drug development

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä	Sanni Virtanen				
Työn nimi	Osakekurssin romahdusriskin ennustaminen lääketeollisuuden yrityksissä – Näyttöä Yhdysvalloista vuosilta 2002-2016				
Tutkinto	Kauppatieteiden maisteri				
Koulutusohjelma	Laskentatoimi				
Työn ohjaaja	Henry Jarva				
Hyväksymisvuosi	2017	**Sivumäärä**	54	**Kieli**	Englanti

Tiivistelmä

Tutkimuksen päätavoitteena on löytää selittäviä tekijöitä osakekurssien romahduksille Yhdysvaltalaisissa lääkealan yrityksissä. Lisäksi tutkimus pyrkii selvittämään empiirisessä analyysissä käytettyjen romahdusriskimittareiden luokittelutarkkuutta.

Tutkimus on luonteeltaan kvantitatiivinen, ja siinä hyödynnetään logistista regressiota sekä usean muuttujan lineaarista regressiota. Tutkimuksessa käytetään kahta akateemisessa kirjallisuudessa käytettyä romahdusriskimittaria sekä luodaan yksi uusi romahdusriskimittari. Aineisto sisältää 1303 havaintoa Yhdysvaltalaisista lääketeollisuuden yrityksistä vuosilta 2002-2016.

Keskeisimpänä löytönä havaittiin, että osakekurssin volatiliteetilla on positiivinen yhteys tulevaan romahdusriskiin, kun taas pääoman tuotolla ja yrityksen koolla taseen loppusumman tai markkina-arvon suhteen on negatiivinen yhteys tulevaan romahdusriskiin lääkealan yrityksissä. Lisäksi havaittiin, että tutkimuksessa käytetyt romahdusriskimittarit antavat toisistaan poikkeavan tuloksen 20-30 prosentille tarkastelujaksoista.

Avainsanat romahdusriski, bioteknologia, lääketeollisuus, lääkkeiden kehitys

Table of Contents

List of tables

List of appendices

1. Introduction

1.1. Background of the study

Stock price crashes cause major concerns for active investors and have a disastrous influence on investor welfare (Ak et al. 2016; Kim & Zhang 2016). Further, stock price crashes are not unusual. Accumulated stock returns are found to be asymmetrically distributed, and one of the indications of this phenomenon is that stocks are more prone to large negative returns than to large positive ones (Campbell & Hentschel 1992; Chen et al. 2001). Hutton et al. (2009) find in their cross-industry study covering the period from 1991 to 2005 that 17.1% of the sample companies experienced at least one crash per year.

In the 21st century, there has been a growing interest in stock price crash risk among academic literature (e.g. Chen et al. 2001; Hutton et al. 2009; Kim et al. 2011a; Kim et al. 2011b; An & Zhang 2013; Kim et al. 2014; Callen & Fang 2015; Ak et al. 2016; Chen et al. 2017). The prior literature focuses on searching cross-industry wide determinants for crash risk. Studies find that firm-specific stock price crash risk is positively associated with accounting opacity (Hutton et al. 2009), corporate tax avoidance (Kim et al. 2011a), short interest (Callen & Fang 2015), real earnings management (Francis et al. 2016), CEO overconfidence (Kim et al. 2016) and earnings smoothing (Chen et al. 2017), and negatively associated with corporate social responsibility (Kim et al. 2014) and accounting conservatism (Kim & Zhang 2016).

While a recent cross-industry study shows that earnings announcements trigger approximately 70% of all stock price crashes (Ak et al. 2016), in biotechnology industry reported earnings have nearly immaterial impact on share prices (Ely et al. 2003; Dedman et al. 2008). Biotechnology industry is a research-intensive sector with long development lead times, extensive capital requirements, and complex intellectual property right matters (Hamill et al. 2013). Product development is in the central role of this highly regulated industry, where all the firms must go through and pass the same development stages defined by the regulator (Girotra et al. 2007). To succeed companies need to continuously develop new products since the treatment of diseases is constantly changing and patent expirations increase the competition and reduce the margins of products (Shortridge 2004). Under U.S. GAAP accounting standards, internal research and development costs should be expensed when incurred pursuant to SFAS 2. Therefore, reported earnings for biotechnology firms may be less informative for investors in respect to firm value compared to less research-intensive industries (Dedman et al. 2008). Despite the complex nature of biotechnology sector that differs from many other industries and the significance of stock price crashes, there is little, if any, empirical evidence of determinants of crash risk in this sector. This paper sheds some light on the issue of how investors can avoid value-destroying stock price crashes in the complex biotechnology industry.

1.2. Research objectives

This study examines determinants of stock price crashes in the U.S. biotechnology sector. There is one main objective and one secondary objectives for this thesis. As the main objective, the thesis aims to analyse whether it is possible to forecast stock price crashes in biotechnology sector using a set of variables, and which of those variables are most prominent to predict future crashes. All the predictors are measured on the basis of information available to the investors before the occurrence of the stock price crash. To my knowledge, this is the first study to investigate crash risk in the biotechnology sector. Further, most of the crash risk studies focus on investigating single determinants of crash risk rather than a set of possible predictors for crash risk. The only exception I am aware of is a recent study of Ak et al. (2016).

The secondary objective of this study is to increase understanding of how well the metrics are classifying stock price crashes. Especially, the aim is to find out how well the metrics are agreeing with each other which sample periods undergo crashes and which do not. Two crash risk metrics introduced by prior literature (the metrics of Chen et al. 2001 and Ak et al. 2016) have remarkably different definitions for a stock price crash. Hence, it is expected that those metrics have differences in classification of crashes. To my knowledge, the prior literature, however, has not considered the accuracy of different crash measures or compared the firm-specific classifications of stock price crashes with different metrics.

Further, an additional goal of this study is to create a new crash risk metric that answers to some of the limitations of existing metrics. The academic literature uses mainly two crash risk metrics (the metrics of Chen et al. 2001 and Hutton et al. 2009), which a recent study of Ak et al. (2016) has criticized. Thus, this thesis aims to provide a new alternative or additional measure that would better identify crashes from large samples.

The study is conducted as quantitative analysis applying both logistic and multiple regression to explore the predictive power of a set of variables to forecast crash risk. The crash metric of Chen et al. (2001) and Ak et al. (2016) are used in this study, as well as an additional crash measure. To identify predictors for crash risk, the study tests the relation between stock price crash metrics and ten explanatory variables, which are chosen to be considered in this study based on the previous literature. The final sample includes 1303 observations covering the period from 2002 to 2016. The sample consists of U.S. biotechnology companies that are in the business of developing new drugs (SIC 2833-2836).

In this study three different crash risk metrics are used, and hence there are three different definition for a stock price crash as well. Chen et al. (2001) define a crash as negative skewness in stock return distribution, whereas Ak et al. (2016) define a stock price crash as an abnormally large and sudden drop in the price of a stock. The new metric created as part of this study defines a stock price crash as a large and sudden drop in the price of a stock that remains at the lower level at least for a week. Existing academic literature uses also some other crash risk metrics but those are left out of the scope of this study.

The study focuses on investigating the ability of limited set of explanatory variable to predict future stock price crashes in the biotechnology sector. Thus, the selected explanatory variables might not be the best predictors for stock price crashes and there might be other determinants for crash risk that are out of the scope of this study. The study is limited to the biotechnology sector that includes both pharmaceutical and pure-play biotechnology companies that are in the business of developing drugs, and hence cannot be directly adapted to other industries. This study contributes to the academic literature by increasing the understanding of crash risk determinants in the biotechnology sector.

1.3. Structure of the thesis

This thesis consists of six chapters. After introduction, the second chapter reviews the literature on stock price crashes. In third chapter the reader is familiarized with the drug development process that is in the central role of biotechnology sector. Further, literature from the biotechnology sector related to market valuations and stock prices is presented. Chapter four introduces the three stock price crash metrics used in this study, describes the data and sample, and presents the independent variables used in the regression models. Research findings are presented in chapter five, started by introducing the descriptive statistics and correlations of variables, and examining crash classification consistency between the three crash metrics. This is followed by presenting the regression analysis results for each crash measure, as well as the results for four sensitivity analyses. Finally, chapter six summarizes and concludes the research, and makes suggestions for future research.

2. Stock price crashes

Stock price crashes are abnormally large and sudden drops in the price of a stock (Ak et al. 2016). In other words, stock price crashes are large, negative, firms-specific and market-adjusted return outliers (Jin and Myers 2006). They cause major concerns for active investors and have a harmful influence on investor welfare (Ak et al. 2016; Kim & Zhang 2016). Over the last decade cross-industry wide determinants of crash risk has been extensively discussed in academic literature (e.g. Hutton et al. 2009; Kim et al. 2011a; Kim et al. 2011b; Kim et al. 2014; Callen & Fang 2015; Ak et al. 2016; Francis et al. 2016; Kim et al. 2016; Kim & Zhang 2016; Chen et al. 2017).

A recent study of Ak et al. (2016) investigates factors causing stock price crashes by analysing news reports simultaneous with large crashes (i.e. top 5% in the distribution of used crash metrics) covering a period from July 2012 to June 2014. They find that 67.9 percent of stock price crashes are caused by earnings announcements. The second most common event causing crashes is earning preannouncement by 9.9%. Hence, according to their study, together 77.8% of crashes are earnings-related. Ak et al. (2016) highlight that, besides the earnings-related causes, the only other major category explaining crashes is 'other firm announcements', accounting for 9.3% of crashes, from which majority are biotechnology firms' bad news related to clinical trials of new drugs. Taking into consideration that the study has a cross-industry sample, such a substantial proportion indicates high importance of clinical trial success in the market valuation of biotechnology companies.

The prior studies find several individual determinants of stock price crashes. In many cases the underlying explanation for stock price crashes is related to managers' intentional information management, especially to bad news hoarding (e.g. Jin & Myers 2006; Hutton et al. 2009; Kim et al. 2011a; Kim et al. 2011b; Callen & Fang 2015; Kim et al. 2016). Managers tend to hide bad news hoping that those would be temporary. Once a certain threshold is reached, this behaviour leads to an abrupt release of an accumulated negative information, resulting in a stock price crash (Jin & Myers 2006; Hutton et al. 2009).

One of the enabling factor for bad news hoarding is tax avoidance, which provides tools and mask for managerial opportunistic behaviour leading to higher crash risk (Kim et al. 2011a). The study covering sample of U.S. firms for the period from 1995 to 2008 provides evidence that tax avoidance has statistically significant and positive association with the future stock price crash risk. Kim et al. (2011a) show, however, that, firms with stronger external monitoring have less pronounced relation between tax avoidance and crash risk.

Weak monitoring of investors is also suggested to increase crash risk and exacerbate bad news hoarding behaviour. An and Zhang (2013) find that transient institutional investors, which have high portfolio turnover and small positions in individual firms, are associated with higher crash risk whereas dedicated institutional investors, which have stable ownerships and large positions in companies, are associated with lower crash risk. The authors propose that weak monitoring of transient investors allow managers to hide bad news, leading to a stock price crash.

In the firms with weak external monitoring by institutional investors or weak internal monitoring by the board, corporate social responsibility (CSR) plays an important role in reducing stock price crash risk (Kim et al. 2014). Further, in firms with less effective corporate governance or lower level of long-term institutional ownership, CSR is negatively associated with crash risk. Kim et al. (2014) suggest that a strong CSR-oriented corporate culture mitigates the managers' bad news hoarding behaviour, resulting in lower stock price crash risk.[1] Chen et al. (2017) contribute to the study finding that bad CRS firms smooth earnings to hide bad news increasing the crash risk of the stock, whereas good CRS firms tend not to smooth earnings.

Chief executive officer (CEO) overconfidence can appear in bad news hoarding behaviour (Kim et al. 2016). The tendency of some managers to overestimate their own excellence and the prospects of positive future outcomes serve as a complementary theory for bad news hording and resulting in stock price crashes. Overconfidence can lead the CEO to proceed with negative NPV projects and hide the bad performance leading to a stock price crash. More conservative accounting policies mitigate the relation between CEO overconfidence and future crash risk, indicating the possibility of investors to recognize bad news earlier.

---

[1] A circumstance worth noticing is that Kim et al. (2014) define a crash as the conditional skewness of return distribution following Chen et al. (2001), rather than an extreme and sudden negative return of a stock. While many studies utilize the skewness measure of Chen et al. (2001) in crash risk research, they often include as well a crash measure of Hutton et al. (2009), which focus on the large, negative returns irrespective of the relation between negative and positive returns. Thus, using only the skewness measure narrows down the definition of a stock price crash.

Kim and Zhang (2016) argue that greater conditional conservatism is related to lower future stock price crash risk. Conditional conservatism refers to the propensity to require less evidence to recognize bad news as losses than to recognize good news as gains (Basu 1997). Chen et al. (2017) find supporting evidence about the negative association between accounting conservatism and crash risk. The findings of Kim and Zhang (2016) hold also to changes in the degree of conditional conservatism that are consistently negatively associated with changes in future crash risk. The authors suggest that conditional conservatism decreases the managers' ability to hide bad news leading to a smaller stock price crash risk.

In an environment with higher information asymmetries (i.e. less information is available to investors), the predictive power of conditional conservatism over crash risk is greater (Kim & Zhang 2016). The predictive power is especially stronger for firms with intensive research and development, high industry concentration and low analyst coverage. Kim et al. (2011a) and Chen et al. (2017) find a consistent outcome considering the mitigating effect of analyst coverage to future crash risk.

Utilizing the bad news hoarding theory, Jin and Myers (2006) and Hutton et al. (2009) show that the lack of information transparency increase future stock price crash risk. Jin and Myers (2006) argue that stock price crashes are more common for firms in opaque countries, referring to countries where information availability to investors is diminished. Similarly, Hutton et al. (2009) find that opacity, measured as accrual-based earnings management, predicts crash risk. Accrual-based earnings management refers to use of judgment in financial reporting altering the levels of discretionary accruals. Thus, using earnings management firms are able to hoard bad news, which can eventually lead to a stock price crash.

Hutton et al. (2009) find in their time-series analysis that the relation between opacity and crash risk has diminished in the post-Sarbanes-Oxley Act (SOX) period. Sarbanes-Oxley Act is a United States federal law, which was enacted in 2002 as a response to number of major corporate and accounting scandals such as Enron Corporation, Tyco International plc and WorldCom. After the passage of SOX, the level of accrual-based earnings management has declined significantly, and at the same time the level of real earnings management activities have increased significantly (Cohen et al. 2008; Francis et al. 2016). Consistent with a survey and field interview study of Graham et al. (2005), the results of Cohen et al. (2008) and Francis et al. (2016) suggest that firms have switched from using accrual-based earnings management to real earnings management. In other words, firms have switched from using accounting actions (i.e. altering the level of discretionary items) to real actions such as delaying maintenance or adjusting R&D expenses to meet short-term earnings benchmarks. Consistently, Francis et al. (2016) show that firm's deviation in real operations from industry norms increases crash risk. The association is especially strong in firms using real operations for earnings manipulation purposes.

A recent study of Chen et al. (2017) shows that earnings smoothing is associated with greater likelihood of stock price crashes, suggesting that earnings smoothing is detrimental to shareholders and can potentially destroy shareholder wealth if the crash risk is realized. The authors suggest that earnings smoothing demonstrates managerial opportunism in financial reporting, especially, through bad news hoarding behaviour. Managers hide poor performance in the hope of better performance in the future. Consistently, earnings smoothing and positive discretionary accruals together are associated with more pronounced crash risk. Further, the authors find that better analyst coverage, higher institutional holdings and good corporate social responsibility of the firm mitigate the stock price crash risk in the presence of earnings smoothing.

Callen and Fang (2015) provide evidence that short interest is positively associated with future crash risk. Short interest describes number of shares sold short in relation to the total number of shares outstanding. The authors suggest that short sellers are able to recognize managers' bad news hoarding behaviour in firms whose stock they short sell in anticipation of stock price crashes. Callen and Fang (2015) find that the association between short selling and crash risk is more pronounced in firms that have weaker external monitoring mechanism, excessive risk-taking behaviour, and higher level of information asymmetry between managers and shareholders.

In the field of crash risk research a well-known paper of Chen et al. (2001) studies crashes as conditional skewness of the stock return distribution. Chen et al. (2001) point out that they are not trying to forecast negative expected returns due to the narrow definition of stock price crashes they have adopted. Yet, the definition has been criticized since skewness of stock return can be either caused by negative long tail effect (i.e. extreme negative returns) or negative fat tail effect (i.e. multiple small negative returns) (Ak et al. 2016). Further, skewness as a measure of crash risk doesn't capture extreme negative outcomes which occur in a period where both extreme negative and extreme positive outcomes are present and therefore returns in the period are normally distributed. Nevertheless, the skewness measure of Chen et al. (2001) is widely used in a crash risk literature (e.g. Kim et al. 2011a; Ak et al. 2016; Francis et al. 2016; Kim & Zhang 2016; Kim et al. 2016; Chen et al. 2017).

The conditional skewness research of Chen et al. (2001) shows that crash risk is more prominent in firms (1) which past returns have been positive over the prior 36 months, (2) that have glamour stocks (i.e. book-to-market ratio of the stocks is low), (3) that have experienced larger trading volume relative to market trend in the prior six months, and (4) that have larger market capitalization. Chen et al. (2001) suggest that larger market capitalization firms are more negatively skewed since they are not able to hide bad news from market as easily as smaller market capitalization firms. This suggestion is contradictory to several crash risk studies utilizing bad news hoarding theory as the explanation for their findings. The studies suggest that crashes occur due to the bad news hoarding behaviour. Managers tend to hide bad news and once certain threshold is reached the accumulated bad news are released to the market leading to a stock price crash. Thus, according to the theory if larger market capitalization firms are not able to hide bad news as easily, those should be less crash prone firms than smaller market capitalization firms.

However, Chen et al. (2001) investigate conditional skewness of returns which might result also from fat left side tail as explained earlier. Hence, the reasoning behind the connection of the market capitalization and skewness can be intuitively accurate if skewness is caused by number of small negative returns. As Chen et al. (2001) argue, managers tend to publish good news immediately and simultaneously hide bad news that might lead to positive skewness before more bad news accumulate and stock price eventually crashes.

The research of Ak et al. (2016) is one of the only studies trying to identify a set of crash predictors instead of identifying single characteristics that forecast stock price crashes. They find five explanatory variables that are associated with future crash risk: abnormally high trading volume, book-to-market ratio, accounting opacity, short interest (referring to the number of shares sold short), and forecasted sales growth. First, abnormally high trading volume has a positive relation with crash risk, indicating that disagreements among investors increase the future crash risk. Second, consistent with Chen et al. (2001), book-to-market ratio is negatively associated with future crash risk.

Third, Ak et al. (2016) find that accounting opacity has positive relation with crash risk, indicating that the use of subjective assumptions in accounting is related to higher likelihood of crashes. The finding is consistent with the study of Hutton et al. (2009), which introduces the connection between accounting opacity and crash risk. Fourth, short interest has positive association with future crash risk. Short sellers are typically experienced investors specialized to distinguish overpriced stocks implying that higher short interest reflects the short sellers' negative view of the stock. Callen and Fang (2015) find consistent connection between short interest and future crash risk. Finally, Ak et al. (2016) show that forecasted sales growth is positively related to the future crash risk, suggesting that the optimistic expectations of sell-side analysts is related to higher likelihood of future crashes.

3. Biotechnology sector

Biotechnology is a collection of techniques that use living cells to develop products such as drugs, food products and chemicals (Hand 2001). Both traditional pharmaceutical firms and pure-play biotechnology firms perform biotechnology research, however, pharmaceutical firms can also develop and manufacture drugs using non-biotech techniques (Ely et al. 2003). In this study, biotechnology term is used to describe both pharmaceutical firms and biotechnology firms, which are in the business of drug development and hence, leave out other areas of biotechnology. Further, in this study biotech and biopharmaceutical terms are used as synonyms for biotechnology. This chapter start by introducing the drug development process in the U.S. followed by reviewing literature about drug development failures and market valuation of stocks in biotechnology sector.

3.1. Drug development process in the U.S.

Drug development and approval process can be divided into five stages.[2] First is discovery and development stage where new molecular compounds are being tested to find new potential drugs. Second stage is called as preclinical research. Drugs go through laboratory and animal testing to examine safety and possible toxicity of the drug. Only one out of 1000 molecular compounds passes the preclinical testing, which takes on average 3.5 years (Dedman et al. 2008).

On the third stage, called as clinical research, drugs are tested on people to ensure safety and effectiveness of the drug. The clinical research is divided into three phases. In phase I, drug is examined with 20-100 volunteers to determine the most frequent side effects. The phase can take several months and approximately 70% of drugs pass the trial. In phase II, up to several hundred patients with the disease or condition participate the trial that takes from several months to two years. The U.S. Food and Drug Administration (FDA) reports that in this phase on average every third drug moves to the next phase. The third and last phase of clinical research include from 300 to 3000 participants who have the disease or condition.

---

[2] The approval process can be found from the web page of FDA:
https://www.fda.gov/ForPatients/Approvals/Drugs/default.htm

The trial takes usually from one to four years, and approximately 25-30% of drugs pass the phase.

After the clinical trials, on fourth stage, a drug developer files an application called New Drug Application (NDA) which must contain all the essential data of the drug. If the application is complete, FDA review team has from 6 to 10 months to decide whether to approve the drug. However, Dedman et al. (2008) argue that in almost all cases FDA approval exceeds that limit, and it takes approximately 2.5 years for FDA to review the NDA. A more recent study of Hamill et al. (2013) reports that the review process takes on average 540 days (≈1.5 years). In the last stage, once a drug gets to the market, the firm is obligated to submit periodic safety updates to FDA in case of unexpected adverse events.

The overall drug approval percentage from clinical trials (i.e. drugs that enter clinical trials and will eventually be approved) reported by FDA seems to be more conservative than the estimates of recent studies. According to FDA, the clinical trial success rate is approximately 6-7%[3] while DiMasi et al. (2016) estimate the rate to be on average 11.83%. An earlier study of DiMasi et al. (2003) estimates the clinical trial success rate to be 21.5%, which is nearly 10 percentage point larger than in the authors' more recent study.

Totally the drug development process from discovery to market approval of a new drug takes approximately 12 years (Dedman et al. 2008), and costs around 2558 million U.S. dollars (DiMasi et al. 2016). The cost represents an average capitalized R&D cost per new approved drug for pharmaceutical companies including the costs of abandoned compounds. The earlier study of DiMasi et al. (2003), using a consistent methodology, has found that the average cost of a new drug was 802 million U.S. dollars in 2000. DiMasi et al. (2016) suggest that the increases in costs are driven by growth in the real out-of-pocket costs of development and by lower success rates for clinical trials. Comparing pharmaceutical companies to pure-play biotechnology companies DiMasi and Grabowski (2007) find that the average development cost and time for a new drug are relatively similar for these two types of companies.

[3] The clinical trial success rate is calculated using approval rates that FDA has reported for each clinical research phase: 70%, 33% and 25-30%. Note that the third phase success rate includes also an approval of FDA review stage during which New Drug Application is either approved or rejected.

3.2. Drug development failures

Drug development failures are negatively affecting the market valuation of the firm and thus, the shareholder wealth (e.g. Sharma & Lacey 2004; Girotra et al. 2007). However, the consequences of drug development failures to a firm's market value are dependent on the stage of the drug development process. In early stages, not only firms are disclosing less information of drug development but also investors are not reacting as strongly to the news as in later stages of the process (Ely et al. 2003; Dedman et al. 2008). Especially later stage announcements that are made after the completion of clinical trial's phase II are causing much stronger market reaction than early stage drug development announcements (Dedman et al. 2008). In the later stage of the drug development process investors begin to treat research and development costs as probable future revenue-generating assets, and hence the declines in share prices are larger when bad news reach the market (Ely et al. 2003).

The event study analysis of 41 New Drug Application rejections by the U.S. Food and Drug Administration shows that negative announcements of drug development outcomes are greatly affecting to the firm's value (Sharma & Lacey 2004). On the event day of product development failure (i.e. the day a firm disclosed the decision of FDA to reject a drug) abnormal returns declined on average 11%, and in a three-day window (from a day before the event to a day after the event) cumulative abnormal returns declined on average 21%. NDA rejections are causing large and abrupt drops in stock prices that, given a broad definition (e.g. Ak et al. 2016), can be called as stock price crashes.

While effects of NDA rejections are massive (21% decline in a 3-day window), Girotra et al. (2007) find that a phase III drug failure causes on average 1.46% decline in stock price during a seven-day event window (from day -2 to day 4). The effect of NDA rejection is particularly strong because it indicates managerial failure in implementing internal controls and procedures to ensure drug safety (Girotra et al. 2007).

Unlike New Drug Application approvals, FDA does not disclose its decision to delay or reject an application (Sharma & Lacey 2004). Thus, companies can decide whether they disclose the bad news about FDA's concerns and rejection decisions. Sharma and Lacey (2004) argue that due to potentially large reputational and legal costs firms are likely to release the bad news related to NDA. On the contrary, Dedman et al. (2008) find that companies are disclosing considerably more good news than bad news even though the probabilities to success in a drug development process is low. From the sample of 234 announcements only twelve were classified as bad news, comprising only 5% of the sample. Considering that only one out of 5000 drugs that undergo preclinical testing is approved to the market (Dedman et al. 2008), the amount of negative news is remarkably low. The existing cross-industry literature of stock price crash risk proposes that bad news hoarding is notable reason for increased crash risk (e.g. Jin & Myers 2006; Hutton et al. 2009; Kim et al. 2011a; Callen & Fang 2015; Kim et al. 2016), which indicates that also in biotechnology industry stockpiling bad news, such as drug development failures, can lead to a stock price crash.

3.3. Market valuation of a stock in the biotechnology sector

While in cross-industry studies earnings announcements are found to be impacting on the market value of a company (e.g. Ak et al. 2016), in the biotechnology sector earnings announcements do not have a significant explanatory power on share prices (Ely et al. 2003; Dedman et al. 2008). In the research-intensive biotechnology sector drug development announcements are impacting much stronger on share prices than earnings announcements or any other type of announcements (Dedman et al. 2008). Consistently, Hamill et al. (2013) state that drug development announcements at various stages are key value drivers for biopharmaceutical firms. Other non-financial measures that are significantly associated with the share price of a firm include for examples the number of drugs a firm has in clinical trials and the number of New Drug Applications submitted to FDA (Ely et al. 3003).

Even though many studies empathize the impact of non-financial indicators on the market value of biotech firms, Ely et al. (2003) argue that financial measures have independent and significant explanatory power over the market value. Their core sample consists of 83 pure-play biotech firms with no approved drugs including 193 firm-year observations over the period from 1988 to 1998. Firms without approved drugs are typically younger and smaller, and their current earnings are not likely to reflect the future potential of the firm. Ely et al. (2003) find that even though earnings are not significantly related to share price, book value and research and development expenses account for 68% of the variation in the market value of a pure-play biotechnology firm. However, the explanatory power of research and development expenses is larger for sample firms with a high-potential portfolio of drugs than for low-potential drug portfolios (82% versus 70%). This is in line with Shortridge's study

(2004) stating that non-financial measures complement the mandatory financial information. Other key financial indicators for biopharmaceutical firms include retained cash, rate of cash expenditure, ‘burn-rate’ (i.e. how quickly a firm uses its shareholder capital), and the timing and source of additional capital requirements (Hamill et al. 2013).

The result of Ely et al. (2003) is also supported by findings of Amir and Lev (1996) who research value-relevance of financial information in telecommunication industry, and conduct supporting analysis from biotechnology industry. They discover that book values are positively associated to share prices while earnings are negatively associated to share prices in biotech firms. Due to the irrational negative relation between earnings and share prices they suggest that in biotechnology sector the reported earnings are inadequate, and provide only little value-relevance to investors.

4. Research design

In this chapter the research design is introduced in three stages. In the first stage, three different metrics are presented to define and identify stock price crashes. Further, limitations of the metrics are discussed. In the second stage, the data and sample are described, and the data modifications are presented. In the last stage, the independent variables are described and the choice of the variables is reasoned.

4.1. Stock price crash metrics

Prior research uses primarily two different stock price crash metrics: skewness measure of Chen et al. (2001) and crash likelihood measure of Hutton et al. (2009).[4] Firstly, Chen et al. (2001) measure crashes in respect of negative skewness of the stock return distribution. The measure for negative coefficient of skewness (NCSKEW) is calculated by dividing the negative of daily return skewness with the standard deviation of daily returns raised to the third power:

$$NCSKEW_{i,t} = -\frac{n(n-1)^{\frac{3}{2}}\sum R_{i,t}^{3}}{(n-1)(n-2)(\sum R_{i,t}^{2})^{\frac{3}{2}}} \tag{1}$$

where $R_{i,t}$ is the sequence of daily market-adjusted returns to stock i during period t, and n is the number of available observations on daily stock returns during the period. However, as this study is focused on pharmaceutical industry and Chen et al. (2001) conduct a cross-industry study, instead of using market-adjusted stock returns, biotechnology industry-adjusted returns are used. Dividing the raw skewness with the cubed standard deviation allows the comparison across stocks with different variances (Chen et al. 2001).

NCSKEW measure shows whether the daily returns of stock i are lognormally distributed during a period t. The minus sign in front of the equation indicates that the larger the value the more crash prone the stock is (Chen et al. 2001). Thus, positive NCSKEW value implies that the left tail of the stock return distribution is either longer or fatter than the right tail of

[4] The existing literature uses also other stock price crash measures in a lesser extent mostly as subsidiary metrics (e.g. Chen et al. 2001; Ak et al. 2016) but those are out of the scope of this study.

the distribution (Ak et al. 2016). The logic behind this measure is that a stock price crash will result in an extreme left tail outcome which can be measured as negative skewness. The recent study of Ak et al. (2016), however, criticizes the NCSKEW measure arguing that it has two limitations. First, NCSKEW misclassifies stocks that have several small negative returns during the period (i.e. a fat left tail) as crashes while according to the definition of a crash only large negative returns (i.e. a long left tail) should be classified as crashes. Second, the measure do not identify crashes that occur in a period that has both crashes and jumps (i.e. large and sudden decreases and increases in stock price) since the return distribution is not skewed.

The second widely used stock price crash measure, established by Hutton et al. (2009), recognizes a crash if at least one weekly return of a stock falls 3.09 standard deviation below its mean value within a fiscal year. The threshold of 3.09 is chosen to produce a frequency of 0.1% in the normal distribution meaning that if returns are normally distributed, 0.1% of the sample firms would be expected to crash in any week. To demonstrate this further, using the average standard deviation of weekly returns across the sample of Hutton et al. (2009), which is 5.8%, abnormal weekly return identified as a crash would be -18% or less.

Ak et al. (2016) argue that the crash measure of Hutton et al. (2009) addresses the two limitations of skewness measure described above but it still has its own limitations. The measure is binary in nature (i.e. it gives only values of 0 or 1) which means that it is not capturing the relative magnitude of the crash. Further, a limitation which applies to both above measures is that they can misclassify crashes which standard deviation of stock returns increases after the crash (Ak et al. 2016). Post-crash return distribution is used to identify crashes unless the crash occurs in the last day (or week in case of the metric of Hutton et al. 2009) of the measured period. To respond these limitations Ak et al. (2016) have conducted a modified version of Hutton et al. (2009) crash measure that is the negative ratio of the period's minimum daily return to the standard deviation of *previous period's* daily returns:

$$CRASH_{i,t} = \frac{-Min(R_{i,t})}{\sqrt{\sum R_{i,t-1}^2/(n-1)}} \tag{2}$$

where $R_{i,t}$ is the sequence of daily market-adjusted returns to stock i during period t, and n is the number of available observations on daily stock returns during the period. In this study, CRASH as presented above is used. However, as with the NCSKEW, biotechnology industry-adjusted returns are used instead of market-adjusted returns.

One of the limitations of CRASH suggested by this study, is that it only takes into consideration the worst return of the period, and do not consider the returns of the following days. For examples, if the share price drops 20% but on the next day it recovers to the same level or even over it, can we consider that as a crash? According to the definition of Ak et al. (2016), a crash is abnormally large and sudden drop in a price of a share. The definition does not emphasize that the stock price should stay in the same level for a while. However, if we think about the shareholder wealth perspective, the one day drop destroys the wealth of shareholders only if they sell the shares on that day.

Using intuition, it can be said that from investor-perspective drops in share prices which are larger than 10 percent are already crashes for shareholders even though the returns would have been volatile in the past as well. Due to the above reasons, I construct an additional crash measure, DROP, which is not affected by the past movements of the stock, and which considers the post-crash returns as well. DROP defines a crash as at least 10% drop in a daily stock return R, where the stock price remains down at least for the five following days. Stock returns need to stay 8% lower for five following days (d+1, d+2, d+3, d+4 and d+5) compared to level at a day before a drop d-1. The figures for each day are calculated as follows:

$$d + 1 = (1 + R_{i,d}) * (1 + R_{i,d+1}) - 1 \tag{3}$$

$$d + n = (1 + R_{i,d}) * (1 + R_{i,d+1}) * \dots * (1 + R_{i,d+n}) - 1 \tag{4}$$

To consider that R is only available for trading days (i.e. not for weekends and holidays) the metric allows that the following five days after a drop can occur, at most, during the next eight days. For this metric, the same daily biotechnology industry-adjusted stock returns are used as with the other two measures.

As in the other measures there are also some limitations related to DROP. This measure, like the original measure of Hutton et al. (2009), is binary in nature which means that the information related to the magnitude of the crash is not utilized, and the thresholds used to identify crashes are arbitrary. There is no single universally accepted exact definition of crash. That's why it can be argued that the threshold of Hutton et al. (2009) is also arbitrary.

Following Chen et al. (2001) and Ak et al. (2016) I compute each of the crash measure using adjusted daily with-dividend stock returns. Since the sample focuses on biotechnology companies, daily returns are adjusted with Fama and French value-weighted drugs industry index[5] which is formed using companies having same SIC codes as the sample of this study (SIC 2833-2836). Adjusting daily returns with industry index is essential since it eliminates crashes that are solely caused by industry-wide declines or broad market declines (Hutton et al. 2009).

Adapting to Chen et al. (2001) the daily industry-adjusted returns are calculated as a log change in stock i less the log change in the Fama and French value-weighted drugs industry index for that day ($R_{i,d} = \ln(1+RET_{i,d}) - \ln(1+Mt_d)$)[6]. Log changes are used since they allow for natural benchmark – for examples NCSKEW gives a value of zero if stock returns are lognormally distributed. Using simple percentage returns instead of log changes would lead to a pronounced correlation between skewness and volatility, and it would cause returns to look more positively skewed in general (Chen et al. 2001).

Expecting that all the firms have published their annual financial statements within four months of the fiscal year end date, the crash measurement period is set to cover eight-month period starting on the following year's May 1 for those firms that are using calendar year as their fiscal year. On the other words, the model is trying to explain stock price crashes during May to December by looking at the figures of previous calendar year.

NCSKEW and CRASH measures are both interpreted in a way that the larger the figure the stronger the crash. When the standard deviation of the returns is not skewed NCSKEW gives a value of zero. Thus, it can be interpreted that if a figure is below zero there are no crashes during the period but it might not be correct to interpret that all the positive values means there is at least one crash. CRASH measure on the other hand gets almost exclusively

[5] Fama and French value-weighted 49 industry portfolios is downloaded from here: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_49_ind_port.html
For examples Hutton et al. (2009) use the Fama and French industry index to adjust the returns.

[6] $RET_{i,d}$ is non-adjusted daily with-dividend return for stock i in a day d, and Mt_d is Fama and French drugs industry index in a day d.

positive values since it's highly unlikely that within an eight-month period daily returns wouldn't be negative at even once.

4.2. Data and sample description

Initially, the sample of this study is combined from CRSP and COMPUSTAT merged annual financial data and CRSP daily stock return data covering a 15-year period of biotechnology industry (SIC 2833-2836) from 2002 to 2016. The sample period is limited to year 2002 in order to ensure the comparability of results between the examined years: SFAS 142 became into effect at the beginning of 2002 stating that amortization of goodwill was no longer permitted and annual impairment testing became mandatory.

Similar to Ak et al. (2016), firm-years having market capitalization less than 100 million at the end of a fiscal year are excluded to ensure that the researched companies are investable. Following Hutton et al. (2009) crash periods with less than half of the stock-return data during the chosen period are excluded, meaning that in a period from May to December there needs to have at least 85 observations[7]. To simplify the data processing, the firm-years which fiscal year is other than calendar year are excluded. Finally, firm fiscal years with insufficient financial data to calculate the variables are excluded. The final sample includes 1303 observations with crash metrics covering eight-month period from a year j and independent variable data from a previous year j-1. Table 1 presents the data modifications in more detail.

Following Ak et al. (2016) and Chen et al. (2017) all continuous variables (i.e. all non-binary crash measures and independent variables) are winsorized at 1st and 99th percentiles to reduce the effect of outliers. This means that all data below the one-percent left tail is replaced by the value of 1st percentile, and all data above the one-percent right tail is set to the value of 99th percentile. On the other words the values of the lowest (highest) one percentage tail is replaced by the next lowest (highest) value to mitigate the distorting effect on statistics.

[7] There are on average 170 trading days in the U.S. in a period from May to December. This is from 245 (days in a period from May to December) x 5/7 (proportion of work days per week) - 5 (holidays) = 170. The holidays are Memorial Day, Independence Day, Labour Day, Thanksgiving Day, and Christmas Day.

Table 1 – Data modifications

	No of cases an action removed	*Size of a sample*
Initial sample of *annual financial data*		5 920 cases
Initial sample of *daily stock return data*		784 922 cases
Excluding cases where:		
1) Stock return data is missing	20 006 cases	764 916 cases
2) Data date is between January and April	246 939 cases	517 977 cases
Aggregating stock return data to eight-month periods from May-Dec based on a year		3 314 cases
Merging annual financial data and annual stock return data (this leaves data where both aggregated stock return data for year j and annual financial data for year j-1 exists)		2 680 cases
Excluding cases where:		
1) Market capitalization is less than 100 million dollars	894 cases	1 786 cases
2) Number of stock return observations in a period is less than 85	75 cases	1 711 cases
3) Firm's fiscal year end month is not a calendar year	230 cases	1 481 cases
4) Insufficient data to count all the needed variables and crash metrics	178 cases	1 303 cases
All continuous variables are winsorized at 1st and 99th percentiles		1 303 cases

Table presents the main data modification steps conducted, starting from two initial samples that are combined to one final sample of 1303 cases.

4.3. Independent variables

Based on the previous literature and the intuitive expectations ten independent variables are chosen to test whether they forecast upcoming stock price crashes or not (Table 2). The independent variables are covering data from a full fiscal year which is in this case also a calendar year. This study uses an assumption that all the firms publish their financial annual reports within four months of a fiscal year end date. Thus, the crash metrics cover an eight-month period from May to December on year j, and independent variables use information from a previous year j-1 that would have been available to investors at the start of each crash period.

Table 2 – Independent variable definitions

Name of the variable	*Explanation*	*Formula*
GOODWILL	Goodwill rate	Goodwill/Total assets
GDWLIP	Impairment of goodwill	Dummy variable
XRD	Research and development expense rate	R&D/Net sales
MTB	Market-to-book ratio	Market Capitalization/Common ordinary equity
logMCAP	Market value of a firm	Log(Market capitalization)
SIZE	Size of a firm	Ln(Total assets)
RCF	Retained cash flow	Change in cash and cash equivalent/Shareholders' equity
VOLATILITY	Volatility of a firm's share price	Standard deviation of daily returns
ROA	Return on assets	Net income/Total assets
LOSS	Net loss	Dummy variable

Table shows definitions for each explanatory variable used in this study. Note that GDWLIP and LOSS are binary variables that are receiving a value of 0 or 1.

Goodwill rate (GOODWILL) is measured as goodwill divided by total assets at the end of the fiscal year. Dividing by total assets allows the comparison between companies of different size. Andreou et al. (2017) find that goodwill has a positive and significant association with future crash risk, indicating that larger goodwill increases the future stock price crash risk.

Since the start of the sample period, year 2002, amortization of goodwill was no longer permitted and impairment testing became mandatory. If the carrying amount of goodwill (i.e. book value) exceeds its fair market value, an impairment loss is recognized. A dummy variable for impairment of goodwill (GDWLIP) is created that equals one if a firm has reported goodwill impairment losses during a fiscal year, and zero otherwise. Missing values of goodwill impairments are replaced with zero.

Ely et al. (2003) find that the market value of companies that focus purely on drug-development, and has no marketable products, is positively associated with both book value and research and development expenses. This study measures research and development expense rate (XRD) calculating the sum of research and development expenses for a fiscal year and dividing it with net sales of the same period. The comparability of the XRD is improved by diving the figure with net sales.

Market-to-book ratio (MTB) is measured as market capitalization (i.e. closing price of a stock for a fiscal year multiplied by number of common shares outstanding) divided by a common ordinary equity which represents the common shareholders' interest in the company. Chen et al. (2001) find that negative skewness is greater in stock with low book-to-market ratio. Thus, it is expected that correspondingly the market-to-book ratio, which is an inverse figure to book-to-market, would have a positive connection with crash risk. Supporting the expectation, Kim et al. (2011a) find a weak positive association between market-to-book value and future crash risk. However, Kim et al. (2016) could not find any relation between crash risk and market-to-book ratio.

The market value of a company (logMCAP) is measured as the common logarithm of market capitalization (i.e. logarithm with base 10). Chen et al. (2001) and Chen et al. (2017) find that larger market capitalization is associated with higher crash risk. Similarly, Kim et al. (2011a) find weak positive association between market value and crash risk. The size of a company (SIZE) is measured as the natural logarithm of total assets. Andreou et al. (2017) find that size of a firm is negatively related to stock price crash risk.

Hamill et al. (2013) state that retained cash and rate of cash expenditures are examples of key performance indicators for biopharmaceutical companies. In this study retained cash flow variable (RCF) is included which is measured as a change in cash and cash equivalents during a fiscal year divided by a shareholders' equity.

The volatility of a company's share price (VOLATILITY) is measured as standard deviation of daily industry-adjusted returns during the previous period t-1. The variable is testing if historical volatility predicts upcoming crashes. While CRASH and NCSKEW standardize the crash results dividing the numerator with standard deviation of daily returns, DROP is an unstandardized crash metric. Hence, it is expected that the crash risk metrics differ in the association of VOLATILITY.

Return on assets (ROA) is measured dividing net income with total assets. ROA is an indicator of how profitable a firm is relative to its total assets. Kim et al. (2011a) and Chen et al. (2017) find a negative association between return on assets and future crash risk.

Finally, a dummy variable of net loss (LOSS) is included to the model. The variable is set equal to one for a firm which made loss during a fiscal year (i.e. net income is negative), and otherwise LOSS is set equal to zero.

5. Research findings

This section presents the research findings in four parts. First, the descriptive statistics and correlations for crash metrics and independent variables are presented. Secondly, the consistency levels between crash metrics are provided, indicating how well the metrics are recognizing crashes from the sample. On third stage, regression analysis results for the three crash measures are presented and discussed. Finally, four sensitivity analyses are introduced to demonstrate how small changes in the methodology affect to the results.

5.1. Descriptive statistics and correlations

Table 3 shows the descriptive statistics for the three stock price crash measures and for ten independent variables used in the analysis. Recall that CRASH and NCKEW are presented such that a higher value denotes a larger crash. The mean value of DROP is 0.429 indicating that 42.9% of sample have at least one crash during a sample period. This is because DROP is binary by nature receiving value 0 if there are no crashes during a period and value 1 if there are at least one crash during a period. The figure is considerably higher than what Hutton et al. (2009) find in their cross-industry study from 1991 to 2005. They report that 17.1% of their sample firm-years experience at least one crash. The difference might indicate that the biotechnology industry experiences more crashes than other industries on average, or the amount of crashes has increased over time, or that DROP identifies more crashes on average than the crash measure of Hutton et al. (2009). While DROP requires at least 10% drop in a daily stock return that must remain at the lower level for a week to be identified as a crash, the measure of Hutton et al. (2009) requires at least 18% drop in weekly return when using the average standard deviation of weekly returns across the sample of Hutton et al. (2009), which is 5.8%, in order to be identified as a crash. Hence, it is probable that DROP identifies more crashes on average than the crash metric of Hutton et al. (2009).

The descriptive statistics of CRASH, excluding the value of 5th percentile (Table 3), are notably higher than the corresponding values reported by Ak et al. (2016). While CRASH in this study has a mean of 5.65, a standard deviation of 6.28, a 5th percentile of 1.43, a median of 3.52, and 95th percentile of 18.55, the figures in the study of Ak et al. (2016) are 3.62, 2.39, 1.45, 2.91, and 8.29, respectively. This is in line with the assumption that the biotechnology industry experiences more crashes than other industries on average, since the other two possible explanations presented above are not valid. First, the sample period of Ak et al. (2016) is from July 2001 to July 2014 which is relatively close to the sample period of this study (i.e. 2002-2016). Thus, the explanation that the amount of crashes would have increased over time is not reasonable. Second, the formula used to calculate CRASH is same than Ak et al. (2016) use in their study. Hence, the differences in measures is not a rational explanation for the higher values in descriptive statistics in case of CRASH metric.

Table 3 – Descriptive statistics on crash metrics and independent variables

	Mean	*Standard deviation*	*5^{th} percentile*	*Median*	*95^{th} percentile*	*N*
Dependent variables						
DROP	0.429	0.495	0.000	0.000	1.000	1303
CRASH	5.654	6.278	1.429	3.520	18.55	1303
NCSKEW	0.246	2.565	-3.098	-0.086	5.610	1303
Independent variables						
GOODWILL	0.070	0.106	0.000	0.000	0.310	1303
GDWLIP	0.023	0.150	0.000	0.000	0.000	1303
XRD	18.64	86.56	0.036	0.420	58.87	1303
MTB	3.965	12.80	-3.372	3.865	17.01	1303
logMCAP	3.036	0.916	2.089	2.730	5.051	1303
SIZE	6.123	2.325	3.534	5.267	10.97	1303
RCF	0.044	0.712	-0.748	0.026	0.912	1303
VOLATILITY	0.032	0.017	0.009	0.030	0.067	1303
ROA	-0.177	0.343	-0.829	-0.078	0.214	1303
LOSS	0.578	0.494	0.000	1.000	1.000	1303

Table shows descriptive statistics on crash metrics and explanatory variables over the sample period 2002-2016.

The mean value of NCSKEW is 0.25 indicating that the sample is weakly and negatively skewed on average. Similar to CRASH, all the figures on descriptive statistics table related to NCSKEW except the 5^{th} percentile are higher than the corresponding figures Ak et al. (2016) reported. While in the sample of this study NCSKEW has a mean of 0.25, a standard deviation of 2.56, a 5^{th} percentile of -3.10, a median of -0.09, and 95^{th} percentile of 5.61, Ak et al. (2016) reported the corresponding figures to be -0.26, 1.53, -2.71, -0.24, and 2.28 respectively. Further, Chen et al. (2001) reported that NCSKEW has a mean of -0.26 and standard deviation of 0.94 measured from a cross-industry sample covering the period from July 1962 to December 1998. Both values are lower than the corresponding figures in this study. This result further reinforces the understanding that stocks of biotechnology industry undergo more crashes than stock in other industries on average.

Recall that independent variables GDWLIP and LOSS are binary by nature. Thus, the mean value of GDWLIP indicates that only 2.3% of firm-years reported goodwill impairments, and mean value of LOSS indicates that 57.8% of firm-years reported net loss for a year. Table 3 presents the descriptive statistics for the independent variables in more detail.

Table 4 presents the correlation matrix including Pearson and Spearman correlations for all the variables used in regression analysis. The Pearson correlations between the three crash metrics are all positive and highly significant. CRASH and NCSKEW are correlated the strongest, having a correlation of 0.69. Ak et al. (2016) find in their cross-industry study a correlation of 0.55 between the corresponding crash metrics. The second strongest correlation in this study is between DROP and CRASH, having a correlation of 0.48. DROP and NCSKEW has a slightly weaker correlation of 0.44.

Measured as Pearson correlation DROP is significantly and positively correlated with research and development costs, volatility of past returns, and the dummy variable of net loss, and significantly and negatively correlated with goodwill, market value of a firm, size of a firm, and return on assets. CRASH is significantly and positively correlated with the dummy variable of net loss, and significantly and negatively correlated with goodwill, market value of a firm, size of a firm, retained cash flow, volatility of past returns, and return on assets. NCSKEW is correlating, at the significance level of 0.05 or lower, negatively with goodwill and return on assets.

Table 4 – Correlation matrix

	DROP	CRASH	NCSKEW	GOODWILL	GDWLIP	XRD	MTB	logMCAP	SIZE	RCF	VOLATILITY	ROA	LOSS
DROP		0.481**	0.441**	-0.211**	-0.040	0.134**	-0.032	-0.361**	-0.371**	-0.021	0.320**	-0.293**	0.264**
		0.000	0.000	0.000	0.149	0.000	0.254	0.000	0.000	0.453	0.000	0.000	0.000
CRASH	0.611**		0.690**	-0.096**	-0.010	0.032	-0.020	-0.070*	-0.097**	-0.055*	-0.106**	-0.152**	0.127**
	0.000		0.000	0.001	0.727	0.249	0.473	0.011	0.000	0.048	0.000	0.000	0.000
NCSKEW	0.465**	0.622**		-0.063*	-0.002	-0.001	0.008	-0.021	-0.047	-0.029	0.051	-0.088**	0.050
	0.000	0.000		0.023	0.936	0.972	0.763	0.446	0.087	0.296	0.063	0.001	0.070
GOODWILL	-0.267**	-0.043	-0.025		0.126**	-0.107**	-0.035	0.478**	0.562**	-0.023	-0.344**	0.299**	-0.330**
	0.000	0.121	0.368		0.000	0.000	0.210	0.000	0.000	0.415	0.000	0.000	0.000
GDWLIP	-0.040	-0.017	0.027	0.145**		-0.033	-0.011	0.078**	0.132**	-0.017	-0.032	0.046	0.007
	0.149	0.546	0.338	0.000		0.237	0.688	0.005	0.000	0.545	0.251	0.099	0.804
XRD	0.268**	0.059*	0.039	-0.563**	-0.131**		-0.015	-0.104**	-0.131**	-0.021	0.132**	-0.308**	0.182**
	0.000	0.033	0.154	0.000	0.000		0.587	0.000	0.000	0.451	0.000	0.000	0.000
MTB	0.037	0.018	0.031	-0.192**	-0.106**	0.158**		0.057*	0.014	0.185**	-0.028	0.105**	-0.038
	0.186	0.514	0.260	0.000	0.000	0.000		0.041	0.611	0.000	0.312	0.000	0.172
logMCAP	-0.343**	-0.016	0.031	0.525**	0.068*	-0.438**	0.076**		0.955**	0.006	-0.619**	0.473**	-0.552**
	0.000	0.566	0.270	0.000	0.014	0.000	0.006		0.000	0.816	0.000	0.000	0.000
SIZE	-0.356**	-0.035	-0.004	0.601**	0.136**	-0.544**	-0.204**	0.879**		0.005	-0.612**	0.531**	-0.576**
	0.000	0.206	0.894	0.000	0.000	0.000	0.000	0.000		0.855	0.000	0.000	0.000
RCF	0.030	-0.012	-0.011	-0.056*	-0.049	0.035	0.078**	0.017	-0.002		0.051	0.067*	-0.023
	0.280	0.669	0.679	0.045	0.078	0.202	0.005	0.547	0.930		0.065	0.015	0.403
VOLATILITY	0.351**	-0.195**	-0.015	-0.493**	-0.053	0.500**	0.047	-0.688**	-0.683**	0.061*		-0.428**	0.458**
	0.000	0.000	0.591	0.000	0.056	0.000	0.087	0.000	0.000	0.028		0.000	0.000
ROA	-0.307**	-0.085**	-0.021	0.451**	0.032	-0.781**	-0.064*	0.528**	0.601**	0.030	-0.518**		-0.731**
	0.000	0.002	0.457	0.000	0.248	0.000	0.021	0.000	0.000	0.276	0.000		0.000
LOSS	0.264**	0.048	-0.005	-0.477**	0.007	0.722**	0.093**	-0.505**	-0.557**	-0.006	0.521**	-0.855**	
	0.000	0.082	0.858	0.000	0.804	0.000	0.001	0.000	0.000	0.831	0.000	0.000	

Correlations are computed from 1303 observations over the sample period 2002-2016. Pearson correlations are above the diagonal and Spearman correlations below the diagonal. P-values appear below correlations. All continuous variables are winsorized at 1st and 99th percentiles. Here * and ** indicate, respectively, 5% and 1% significance (two-tailed).

5.2. Consistency of crash measures

Since DROP measure is binary by nature and for other two measures there are no predefined limits which would identify crashes from no-crashes, limits are set for CRASH and NCSKEW to compare the results between the measures. The limits are set based on the percentage of crashes DROP identifies. The percentage of eight-month periods during which one or more crashes occur is 42.9 % for each measure. Thus, a period is defined to have at least one crash if CRASH exceeds 3.997, or NCSKEW exceeds 0.096. Note that CRASH and NCSKEW are set as binary variable only under this subsection, particularly in Table 5, and are later preferred as continuous variables unless otherwise stated.

Table 5 – Crash measure consistency matrices

		CRASH		
		One or more crashes	*No crashes*	*Cohen's Kappa*
DROP	*One or more crashes*	32.0 %	10.9 %	0.555**
	No crashes	10.9 %	46.2 %	

		NCSKEW		
		One or more crashes	*No crashes*	*Cohen's Kappa*
CRASH	*One or more crashes*	31.2 %	11.7 %	0.524**
	No crashes	11.7 %	45.4 %	

		DROP		
		One or more crashes	*No crashes*	*Cohen's Kappa*
NCSKEW	*One or more crashes*	28.9 %	14.0 %	0.427**
	No crashes	14.0 %	43.1 %	

Table shows consistency level matrices and values of Cohen's Kappa between the crash metrics. Here ** indicates 1% significance (two-tailed). Note that CRASH and NCSKEW are set as binary variables to allow the comparison between the metrics.

The level of total consistency between the binary metrics is from 72.0 to 78.2 percent meaning that two metrics get the same firm-specific result (i.e. at least one crash or no crashes) with the probability of 72.0-78.2 percent (Table 5). On the other hand, the metrics give a different result for 21.8-28.0 percent of the sample periods. The results of DROP and CRASH measures are the closest from each other, agreeing that the same 32.0% of the periods have at least one crash, and the same 46.2% of the periods do not have any crashes, summing up to the total consistency of 78.2%. This leaves 21.8% of the periods that have conflicting results; the other metric states that there is at least one crash and the other states that there are no crashes during the period.

The second highest consistency level is between CRASH and NCSKEW having a same result for 76.6% of the observations, and the lowest consistency is between DROP and NCSKEW having a same result for 72.0% of the sample periods. You can find the matrices from Table 5 where the results are presented in more detail.

Note that since the amount of crashes is forced to equal despite which metric is being used, the percentages of cases where metrics are disagreeing with each other are the same within a matrix. For examples DROP and CRASH both identify that 42.9% of the periods have at least one crash, but they are agreeing which periods have crashes only with 32.0% of the whole sample. Thus, for 10.9 percentage point DROP identifies at least one crash while CRASH do not, and the same goes for other way around as well.

Cohen’s Kappa presented in Table 5 is a measure of agreement used to determine how well two categorical metrics are reporting same results. The test receives values between 0 and 1, where 1 indicates a perfect agreement between the measures and 0 indicates that the agreement is as good as by chance. Cohen’s Kappa controls for chance factor, meaning that it takes into consideration how good the level of consistency would have been by chance. Thus, each of the values is smaller than the corresponding level of consistency presented above, which is simply calculated based on frequencies. (Watson & Petrie 2010)

The values of Cohen’s Kappa between the three dummy variables are all statistically highly significant. The highest Cohen’s Kappa is between DROP and CRASH, having a value of 0.56. CRASH and NCSKEW have almost as high level of agreement, which is 0.52, whereas

NCSKEW and DROP have the lowest level of agreement, having Cohen's Kappa of 0.43.[8] The strength of agreements between the crash metrics are in the moderate level (Landis & Koch 1977).

5.3. Regression analysis results

In this subsection, the regression analysis results for all the three crash measures, DROP, CRASH, and NCSKEW, are presented. Finally, the regression results are concluded and discussed.

5.3.1. DROP

The study sample consists of 1303 observation periods, each covering an eight-month of stock return data and financial data from a full fiscal year. DROP identifies 559 periods during which at least one crash occurs. By predicting that there are no crashes (i.e. using a null model), the classification accuracy would be 57.1%. The logistic regression model, which results are reported on Table 6, can detect correctly 56.9% of cases when there is at least one crash, and 71.9% of cases when there are no crashes. Hence, the overall predictive capacity of the model is 65.5%, increasing the predictive capacity compared to the null model by 8.6 percentage points.

The logistic regression model includes three statistically significant independent variables: VOLATILITY, ROA, and XRD. First, higher VOLATILITY is associated with higher crash risk, but the magnitude alters notably within the sample. The odds-ratio of VOLATILITY is approximately 5.39 million, indicating that for each one unit increase in VOLATILITY, the odds of having at least one crash during a period, when controlling for other variables, increases around 5.39 million times higher. Note that the range of VOLATILITY in the sample is, however, only 0.09. Thus, it is more reasonable to state that 0.01 unit increase in VOLATILITY leads to approximately 53.9 thousand times greater likelihood that at least

[8] The Pearson correlations between the three dummy crash metrics equal to the values of Cohen's Kappa between the variables. Note that Pearson correlations between DROP, and continuous CRASH and NCSKEW differ from the correlations between the dummy variables (see Table 4).

one crash occurs during a period.[9] The 95% confidence interval for the odds-ratio of VOLATILITY is from 742.6 to 3.9×10^{10} indicating that 95% of odds-ratios is between that range. Hence, firms with high historical volatility in stock returns experience more crashes on average.

Table 6 – Logistic regression results using DROP as the dependent variable

Variable	*Beta coefficient*	*Standard error*	*Wald*	*p-value*	*Odds-ratio*
Intercept	1.058*	0.422	6.269	0.012	2.880
GOODWILL	-0.103	0.752	0.019	0.891	0.902
GDWLIP	0.000	0.452	0.000	0.999	1.000
XRD	0.002*	0.001	4.354	0.037	1.002
MTB	-0.001	0.005	0.085	0.771	0.999
logMCAP	-0.311	0.241	1.657	0.198	0.733
SIZE	-0.172	0.102	2.851	0.091	0.842
RCF	-0.049	0.085	0.326	0.568	0.953
VOLATILITY	15.501**	4.536	11.677	0.001	5393454.131
ROA	-0.658*	0.284	5.368	0.021	0.518
LOSS	-0.122	0.192	0.404	0.525	0.885
Nagelkerke R-Square		0.224			

Table shows regression analysis results for the model using DROP as the dependent variables. Here * and ** indicate, respectively, 5% and 1% significance (two-tailed).

Secondly, higher ROA is related to lower likelihood of crashes. The odds-ratio of ROA is 0.518 indicating that one unit increase in ROA is associated with the 48.2% reduction in the probability of having one or more crashes during the following period. The confidence interval with a 95% confidence level is between 0.297 and 0.904 referring that the reduction in crash risk is typically from 9.6 to 70.3 percent. Thus, firms using more effectively their assets to generate earnings are more likely to experience less stock price crashes.

[9] Due to the abnormally high odds-ratio of VOLATILITY, as a robustness check, the regression is run without the variable. The explanatory power of the model slightly decreases and the odds-ratios of other independent variables do not change substantially, indicating that the odds-ratio of VOLATILITY is distorted, and its magnitude to the likelihood of future crashes is not as strong as the ratio suggests.

Finally, higher XRD is associated with higher crash risk. XRD has an odds-ratio of 1.002 indicating that for one unit increase the likelihood of having at least one crash during a period is 1.002 times higher controlling for other variables. The 95% confidence interval for the odds-ratio of XRD is from 1.000 to 1.003. The results show that firms with larger relative size of research and development costs have slightly higher crash risk on average. Note that since the standard deviations of the three independent variables are varying a lot (i.e. from 0.017 to 86.56; see Table 3) it is not reasonable to compare the magnitude of odds-ratios in isolation.

The model has Nagelkerke R-Square value of 0.224 (Table 6). The measure is rather similar than R-Square used in multiple regression analysis. R-Square is the proportion of total sample variation in the dependent variable that is explained by the independent variable. On the other words, R-Square indicates how well the model is explaining the variation in the dependent variable. Thus, the logistic regression model is explaining roughly 22.4% of the variation in DROP. However, Nagelkerke R-Square is one of the pseudo R-Square measures which are created for logistic regression analysis to approximate R-Square, and for that reason the values of Nagelkerke R-Square and R-Square or Adjusted R-Square cannot be directly compared.

5.3.2. CRASH

Table 7 represents the results of multiple regression analysis using CRASH as the dependent variable. The regression has two statistically significant independent variables, VOLATILITY and ROA, both being highly significant at the 1% level. Both variables have negative association with CRASH referring that higher VOLATILITY and ROA are related to lower crash risk. The unstandardized beta coefficient of VOLATILITY is -106.3 indicating that one unit increase in VOLATILITY is associated with -106.3 unit decrease in the dependent variable CRASH. The 95% confidence interval for the unstandardized beta coefficient is between -131.1 and -81.4, referring that there is a 95% likelihood that unstandardized beta coefficients of VOLATILITY are between the range.

The regression results related to VOLATILITY contradict with the results when using DROP as the dependent variable. I suggest that the conflicting results are caused by the fundamental difference between the two metrics. CRASH normalizes the stock price returns by dividing them with a standard deviation of returns from the last period. Hence, when

VOLATILITY (measured as the last period's standard deviation of daily returns) is high, CRASH interprets that there are no crashes while DROP considers only the magnitude of the absolute drops in the stock returns. Further, I propose that VOLATILITY is negatively related to CRASH since, due to the characteristics of the metric, crashes are only possible if VOLATILITY is low.

Table 7 – Multiple regression results using CRASH as the dependent variable

Variable	*Beta coefficient*	*Standard error*	*t-value*	*p-value*
Intercept	10.356**	1.129	9.169	0.000
GOODWILL	-3.437	1.971	-1.744	0.081
GDWLIP	0.597	1.148	0.520	0.603
XRD	-0.001	0.002	-0.347	0.729
MTB	-0.004	0.014	-0.307	0.759
logMCAP	0.418	0.669	0.625	0.532
SIZE	-0.524	0.285	-1.840	0.066
RCF	-0.246	0.241	-1.023	0.306
VOLATILITY	-106.274**	12.657	-8.397	0.000
ROA	-2.586**	0.779	-3.321	0.001
LOSS	0.784	0.534	1.468	0.142
Adjusted R-Square	0.075			
F-value	11.497**			

Table shows regression analysis results for the model using CRASH as the dependent variable. Here * and ** indicate, respectively, 5% and 1% significance (two-tailed). Beta coefficients are unstandardized.

The unstandardized beta coefficient for ROA is -2.6, indicating that for each unit increase CRASH is going to be on average 2.6 units lower. The confidence interval for beta coefficient in the level of 95% is from -4.1 to -1.1. The regression results related to ROA are in line with the results from logistic regression analysis where DROP is the dependent variable. However, there seems to be no association between CRASH and XRD, unlike with DROP and XRD, since the p-value of 0.73 shows that the relation is statistically highly insignificant.

The value of Adjusted R-Square is 0.075 indicating that 7.5% of the variance in dependent variable is explained by the independent variables. On the other words, the overall explanatory power of the regression is 7.5%. F-value and t-values are indicators of statistical significance where F-value reflects the overall significance of the model whereas t-values reflect the significance of individual predictors. Larger F-value indicates higher significance of the model, and higher t-value in case of positive beta coefficient and lower t-value in case of negative beta coefficient indicate higher significance of the variable. F-value of this model is statistically highly significant, indicating that the explanatory variables in this model are able to explain the variance in the crash risk metric.

5.3.3. NCSKEW

Results of the multiple regression analysis where NCSKEW is the dependent variable are presented in Table 8. The regression has two statistically significant independent variables at the level of 5%: logMCAP and ROA. First, higher logMCAP is associated with higher crash risk. The unstandardized beta coefficient of logMCAP is 0.67, and the 95% confidence interval for the value is between 0.11 and 1.22. Thus, firms with higher market capitalization (i.e. the market value of a firm's outstanding shares) have higher crash risk measured as the skewness of the return distribution. DROP and CRASH do not show similar relation with market capitalization and crash risk.

Secondly, higher ROA is associated with lower likelihood of crashes. ROA has an unstandardized beta coefficient of -0.74 with rather large fluctuation. The 95% confidence interval for beta coefficient is from -1.39 to -0.09. The result is consistent with the results of models using DROP or CRASH as the dependent variable.

Adjusted R-Square value for the model is 0.009, indicating that the model is explaining 0.9% of the variance in the dependent variable. Thus, the explanatory power of the model with NCSKEW as the dependent variable is relatively low. In comparison, the explanatory power of CRASH regression model is 7.5%. The statistical significance of the model, measured with F-value, is 2.2. That is notably lower than in the regression which uses CRASH as the dependent variable, having the F-value of 11.5 (see Table 7). The F-value in this model is statistically significant, referring that the independent variables are, however, able to explain the variance in the dependent variable.

Table 8 – Multiple regression results using NCSKEW as the dependent variable

Variable	*Beta coefficient*	*Standard error*	*t-value*	*p-value*
Intercept	-0.593	0.477	-1.243	0.214
GOODWILL	-0.862	0.833	-1.035	0.301
GDWLIP	0.250	0.485	0.515	0.606
XRD	-0.001	0.001	-1.256	0.209
MTB	0.002	0.006	0.416	0.678
logMCAP	0.668*	0.283	2.362	0.018
SIZE	-0.222	0.120	-1.843	0.066
RCF	-0.105	0.102	-1.036	0.300
VOLATILITY	6.317	5.349	1.181	0.238
ROA	-0.740*	0.329	-2.248	0.025
LOSS	-0.161	0.226	-0.714	0.475
Adjusted R-Square	0.009			
F-value	2.237*			

Table shows regression analysis results for the model using NCSKEW as the dependent variable. Here * and ** indicate, respectively, 5% and 1% significance (two-tailed). Beta coefficients are unstandardized.

5.3.4. Interpretation of the results

Table 9 summarizes the regression results of each model. Four of the explanatory variables are statistically significant in at least one model using one of the three crash measures as the dependent variable. These four variables are ROA, VOLATILITY, XRD, and logMCAP. The most obvious indicator seems to be ROA, having consistently negative association with crash risk, and being statistically significant predictor in each model. Thus, I find that ROA is negatively associated with future crash risk, indicating that more profitable firms in relation to their total assets are less likely to experience stock price crashes. This result is consistent with the cross-industry studies of Kim et al. (2011a) and Chen et al. (2017).

Table 9 – Comparison of regression results between the three main models

	DROP		*CRASH*		*NCSKEW*	
	Beta	*p-value*	*Beta*	*p-value*	*Beta*	*p-value*
Intercept	**1.058**	**0.012**	**10.356**	**0.000**	-0.593	0.214
GOODWILL	-0.103	0.891	-3.437	0.081	-0.862	0.301
GDWLIP	0.000	0.999	0.597	0.603	0.250	0.606
XRD	**0.002**	**0.037**	-0.001	0.729	-0.001	0.209
MTB	-0.001	0.771	-0.004	0.759	0.002	0.678
logMCAP	-0.311	0.198	0.418	0.532	**0.668**	**0.018**
SIZE	-0.172	0.091	-0.524	0.066	-0.222	0.066
RCF	-0.049	0.568	-0.246	0.306	-0.105	0.300
VOLATILITY	**15.501**	**0.001**	**-106.274**	**0.000**	6.317	0.238
ROA	**-0.658**	**0.021**	**-2.586**	**0.001**	**-0.740**	**0.025**
LOSS	-0.122	0.525	0.784	0.142	-0.161	0.475
Nagelkerke/Adjusted R-Square		0.224		0.075		0.009

Table shows the unstandardized beta coefficients and p-values for the variables in each model using different dependent variables.
The bolded values are statistically significant (p-value < 0.05).

Second notable indicator for crash risk is VOLATILITY, having positive and significant association with the dependent variable DROP. On the contrary, VOLATILITY is negatively associated with CRASH and has statistically insignificant relation with NCSKEW. Recall that the crash risk metric DROP is demonstrating the extreme negative stock returns without normalizing the results with standard deviation, whereas CRASH normalizes the most negative stock return of the period with the previous period's standard deviation of the return distribution, and NCSKEW normalizes the skewness of the return distribution with the same period's standard deviation of the returns. Hence, CRASH do not define large negative returns as stock price crashes if the volatility of *the previous period* has been high, and NCSKEW do not define negative skewness of return distribution as stock price crashes if the volatility of *the same period* has been high. On the other hand, VOLATILITY is demonstrating *the previous period's* standard deviation of daily returns. This explain the negative association between VOLATILY and CRASH: if VOLATILITY is high, CRASH do not identify stock price crashes due to the normalized returns. Consistently, the statistically insignificant relation between VOLATILITY and NCSKEW can be justified. As mentioned, NCSKEW is normalized with *the same period's* standard deviation. However, standard deviation of the same period and previous period are intuitively rarely the same, indicating that high VOLATILITY do not automatically refer to the absence of crashes, as it does with CRASH, nor to the higher likelihood of crashes, as with DROP.

Thirdly, XRD has a week positive and statistically significant association with DROP. Ely et al. (2003) find that higher research and development expenses are related to higher market value of a biotechnology firm. I suggest that once the development of a drug is reached to the stage where investors start to treat research and development expenses as probable future revenue-generating assets (Ely et al. 2003), the failure of the drug development might cause a stock price crash. However, in this study the other two crash measures, besides DROP, show that the relation between XRD and crash risk is not statistically significant. Due to the weak association, I suggest that the connection between XRD and crash risk requires further research.

Finally, logMCAP has positive and statistically significant association with NCSKEW, indicating that larger firms in terms of market capitalization are more negatively skewed. Chen et al. (2001) and Chen et al. (2017) find consistent positive relation between market capitalization and the stock distribution skewness measure. The other two crash metrics in the study did not show statistically significant relation between the crash risk and logMCAP. However, results of collinearity statistics show signs of multicollinearity between logMCAP and SIZE, indicating that further analyses are needed to confirm the association between logMCAP and crash risk as well as SIZE and crash risk. The concept of multicollinearity is further explained and the needed analysis are conducted as part of sensitivity analysis in the next subsection.

5.4. Sensitivity analyses

In this subsection, four sensitivity analyses are conducted to test the robustness and sensitivity of the model. Small changes are done to the main regression model related to multicollinearity, winsorized outliers, the length of the crash period, and subsectors of biotechnology industry.

5.4.1. Effects of multicollinearity

As part of the multiple regression analysis collinearity statistics, especially tolerance and variance inflation factor (VIF) are considered, to recognize possible multicollinearity. Multicollinearity refers to a phenomenon where two or more independent variables are highly linearly associated. Even though multicollinearity does not reduce the significance or explanatory power of the model, it may affect to the validity of individual explanatory variable results. The collinearity statistics show that SIZE has a tolerance of 0.06 and VIF of 15.64, and logMCAP has, respectively, values of 0.08 and 13.41, and in addition the correlation between the two variables is 0.96. These values can be considered as significant signs of multicollinearity. The other independent variables are significantly different from SIZE and logMCAP in respect of collinearity statistics, having a tolerance of 0.39 or larger and VIF of 2.6 or smaller, indicating that there is no suspicion of multicollinearity related to the other variables.

The regression analysis is run again for all the three dependent variables leaving first out SIZE, and then leaving out logMCAP, while keeping everything else unchanged, to see how the multicollinearity is affecting to the regression results. Table 10 presents the results for each crash metrics and for each combination of independent variables, including the main models, to allow the comparison.

First, by leaving SIZE out of the model logMCAP becomes statistically highly significant for the models using DROP or CRASH as the dependent variable. Additionally, for the regression using CRASH as the dependent variable GOODWILL becomes statistically highly significant. On the contrary, when SIZE is removed from the model using NCSKEW as the dependent variable the results of logMCAP transform from significant to insignificant.

Secondly, by leaving logMCAP out of the model, SIZE transforms from insignificant to highly significant predictor for the models which are using either DROP or CRASH as the dependent variable. The results of the model using the third metric, NCSKEW, do not experience remarkable changes besides that the number of statistically significant predictors decrease by one when logMCAP (which is statistically significant in the main model) is removed from the model.

Table 10 – Sensitivity analysis results: effects of multicollinearity

	Beta coefficients and p-values								
	DROP			*CRASH*			*NCSKEW*		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	**1.058**	**1.114**	**0.794**	**10.356**	**10.596**	**10.686**	-0.593	-0.492	-0.066
	0.012	**0.008**	**0.031**	**0.000**	**0.000**	**0.000**	0.214	0.300	0.875
GOODWILL	-0.103	-0.541	0.078	-3.437	**-4.811**	-3.727	-0.862	-1.444	-1.325
	0.891	0.441	0.917	0.081	**0.009**	0.052	0.301	0.062	0.102
GDWLIP	0.000	-0.150	0.088	0.597	0.229	0.497	0.250	0.094	0.091
	0.999	0.738	0.842	0.603	0.840	0.662	0.606	0.844	0.850
XRD	**0.002**	**0.002**	**0.002**	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	0.037	**0.037**	**0.039**	0.729	0.723	0.737	0.209	0.207	0.225
MTB	-0.001	0.000	-0.002	-0.004	0.000	-0.003	0.002	0.004	0.005
	0.771	0.973	0.618	0.759	0.979	0.834	0.678	0.481	0.422
logMCAP	-0.311	**-0.674**	X	0.418	**-0.711**	X	**0.668**	0.190	X
	0.198	**0.000**	X	0.532	**0.008**	X	**0.018**	0.092	X
SIZE	-0.172	X	**-0.289**	-0.524	X	**-0.361**	-0.222	X	0.039
	0.091	X	**0.000**	0.066	X	**0.002**	0.066	X	0.418
RCF	-0.049	-0.051	-0.048	-0.246	-0.254	-0.247	-0.105	-0.109	-0.107
	0.568	0.545	0.571	0.306	0.292	0.305	0.300	0.286	0.295
VOLATILITY	**15.501**	**15.695**	**16.483**	**-106.3**	**-105.8**	**-107.5**	6.317	6.508	4.291
	0.001	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	0.238	0.224	0.417
ROA	**-0.658**	**-0.769**	**-0.596**	**-2.586**	**-2.922**	**-2.683**	**-0.740**	**-0.882**	**-0.896**
	0.021	**0.005**	**0.034**	**0.001**	**0.000**	**0.000**	**0.025**	**0.006**	**0.006**
LOSS	-0.122	-0.123	-0.103	0.784	0.776	0.749	-0.161	-0.164	-0.217
	0.525	0.521	0.592	0.142	0.147	0.159	0.475	0.467	0.335
Nagelkerke/ Adjusted R-Square	0.224	0.222	0.223	0.075	0.073	0.075	0.009	0.008	0.006

Table shows the unstandardized beta coefficients and p-values for the variables in each model using different dependent variables and different combinations of independent variables. P-values are presented below the beta coefficients.
(1). The main model with all the ten variables
(2). A model without SIZE
(3). A model without logMCAP
The bolded values are statistically significant (p-value < 0.05).

As a conclusion, when taking multicollinearity into consideration, both SIZE and logMCAP are statistically significant predictors. The variables have a negative association with crash risk, indicating that smaller firms in terms of market capitalization or total assets have higher stock price crash risk than larger firms. However, when defining a crash as negative skewness, there seems to be no connection between the size of the firms or market capitalization of the firms and future crash risk.

Consistently with this study, Andreou et al. (2017) report that smaller firms in term of total assets have higher likelihood to experience stock price crashes in the future. On the contrary to the results of this study, cross-industry studies of Chen et al. (2001) and Chen et al. (2017) report a positive association with market capitalization and crash risk. I suggest that in the biotechnology sector larger firms in terms of market capitalization or total assets are experiencing less crashes since larger firms are not as dependent on the success of one drug as smaller biotechnology firms are. Supporting this suggestion Girotra et al. (2007) find that the failure of phase III clinical trials impacts less to the market valuation of the firms if the firm is developing other drugs for the same market as the failed drug or if the firm is developing other drugs that require same resources as the failed drug used. Intuitively it can be expected that larger firms have more ongoing projects for the same market or more ongoing project that need same resources as the failed drug, and hence larger biotechnology firms are less prone to stock price crashes.

5.4.2. Dealing with outliers

Following the existing academic literature from the field (e.g. Ak et al. 2016; Chen et al. 2017), all the continuous variables are winsorized at 1 and 99 percentiles to mitigate the effect of outliers. As a sensitivity check, the effects are tested by comparing the results of the main model to results using non-winsorized data. I find that using non-winsorized data causes small changes to the results.

Table 11 presents the regression results for each model including the results of the main models. First, the explanatory power of the model when using DROP as the dependent variable stays the same. However, one of the independent variables, XRD, transforms from significant to insignificant, having a p-value of 0.067, which is slightly over the threshold.

Secondly, the predictive power of the model using CRASH at the crash metric is weaker, and there are some differences in the significance levels of predictors when the data is not winsorized. The predictive power measured as Adjusted R-Square reduces from 7.5% to 5.8%, and F-value, which is measuring the significance of the model, declines from 11.5 to 9.0. The independent variables SIZE and LOSS transform from insignificant to significant predictors where SIZE has a negative association and LOSS has a positive association with crash risk. Further, ROA transforms from significant to insignificant predictor.

Finally, the changes with the model utilizing NCSKEW are similar to changes with the model utilizing CRASH. The predictive power declines from 0.9% to 0.4%, F-value reduces from 2.2 to 1.6, and SIZE becomes significant and ROA becomes insignificant predictor.

Table 11 – Sensitivity analysis results: dealing with outliers

	Beta coefficients and p-values					
	DROP		*CRASH*		*NCSKEW*	
	(1)	(2)	(1)	(2)	(1)	(2)
Intercept	**1.058** **0.012**	**1.149** **0.005**	**10.356** **0.000**	**9.717** **0.000**	-0.593 0.214	-0.772 0.106
GOODWILL	-0.103 0.891	-0.078 0.917	-3.437 0.081	-3.283 0.130	-0.862 0.301	-0.706 0.410
GDWLIP	0.000 0.999	-0.012 0.979	0.597 0.603	0.510 0.399	0.250 0.606	0.236 0.641
XRD	**0.002** **0.037**	0.001 0.067	-0.001 0.729	$1.5\text{x}10^{-5}$ 0.949	-0.001 0.209	$-1.3\text{x}10^{-5}$ 0.886
MTB	-0.001 0.771	0.001 0.730	-0.004 0.759	-0.003 0.734	0.002 0.678	0.001 0.848
logMCAP	-0.311 0.198	-0.323 0.172	0.418 0.532	0.733 0.312	**0.668** **0.018**	**0.758** **0.008**
SIZE	-0.172 0.091	-0.178 0.076	-0.524 0.066	**-0.664** **0.031**	-0.222 0.066	**-0.266** **0.029**
RCF	-0.049 0.568	-0.015 0.492	-0.246 0.306	-0.054 0.438	-0.105 0.300	-0.004 0.885
VOLATILITY	**15.501** **0.001**	**14.163** **0.001**	**-106.3** **0.000**	**-94.15** **0.000**	6.317 0.238	8.873 0.076
ROA	**-0.658** **0.021**	**-0.558** **0.029**	**-2.586** **0.001**	-0.264 0.308	**-0.740** **0.025**	0.044 0.669
LOSS	-0.122 0.525	-0.074 0.692	0.784 0.142	**1.888** **0.000**	-0.161 0.475	0.176 0.355
Nagelkerke/Adjusted R-Square	0.224	0.225	0.075	0.058	0.009	0.004

Table shows the unstandardized beta coefficients and p-values for the variables in each model using different dependent variables. P-values are presented below the beta coefficients.
(1). The main model with winsorized variables
(2). A model with non-winsorized variables
The bolded values are statistically significant (p-value < 0.05).

Overall, the test shows that the outliers are affecting to the results at some level. The largest effects are the increase in the significance of SIZE and LOSS, and the decrease in the significance of XRD and ROA. I suggest, however, that winsorizing outliers provides likely more useful information to determine the indicators of crash risk.

5.4.3. Extending the crash period

In the main model, the crash metrics are measured from the beginning of May to the end of December. This leaves a four-month lag between the predictor data and crash risk data. As a sensitivity analysis, I test how extending the crash period to last from January to December affects to the regression results. Due to the change in the crash period, the financial information used as predictors is not available to the investors during the first months of the crash period until the firm publish the annual financial report. Regression results are presented in Table 12, and the main differences between the two models for each crash measure are discussed below.

Extending the crash period improves explanatory power of all the three models having different crash metrics. First, the classification accuracy of the model using DROP as the dependent variable improves 17.9 percentage points compared to the null model (from 52.5% to 70.4%), whereas with the shorter crash period the increase in accuracy is only 8.6 percentage points. Further, the explanatory power of predictors, measured as Nagelkerke R-Square, increases from 22.4% to 30.4% compared to the main model, and XRD transforms from significant to insignificant independent variable.

Secondly, the explanatory power of the model, using CRASH as the dependent variable, increases from 7.5% to 8.7% when the crash period is extended. The significance of the model, measured as F-value, increases slightly from 11.5 to 13.4. Further, GOODWILL becomes a significant predictor, having beta coefficient of -4.9, indicating a negative association between the predictor and crash risk.

The explanatory power of the third model, having NCSKEW as the crash metric, improves from 0.9% to 2.0% when using extended crash period. F-value, the measure for the significance of the model, increases from 2.2 to 3.7. Extending the crash period does not affect notably to the significance levels of the independent variables.

Table 12 – Sensitivity analysis results: extending the crash period

	Beta coefficients and p-values					
	DROP		*CRASH*		*NCSKEW*	
	(1)	(2)	(1)	(2)	(1)	(2)
Intercept	**1.058**	1.297	**10.356**	**11.287**	-0.593	-0.585
	0.012	0.003	**0.000**	**0.000**	0.214	0.280
GOODWILL	-0.103	-1.109	-3.437	**-4.860**	-0.862	-1.640
	0.891	0.139	0.081	**0.023**	0.301	0.082
GDWLIP	0.000	-0.394	0.597	0.123	0.250	-0.055
	0.999	0.403	0.603	0.923	0.606	0.922
XRD	**0.002**	0.002	-0.001	0.002	-0.001	-0.001
	0.037	0.077	0.729	0.412	0.209	0.402
MTB	-0.001	-0.010	-0.004	-0.016	0.002	0.000
	0.771	0.085	0.759	0.286	0.678	0.972
logMCAP	-0.311	-0.316	0.418	0.328	**0.668**	**0.680**
	0.198	0.206	0.532	0.652	**0.018**	**0.034**
SIZE	-0.172	-0.159	-0.524	-0.468	-0.222	-0.226
	0.091	0.135	0.066	0.133	0.066	0.100
RCF	-0.049	0.081	-0.246	-0.091	-0.105	-0.092
	0.568	0.395	0.306	-0.347	0.300	0.425
VOLATILITY	**15.501**	**20.077**	**-106.3**	**-111.2**	6.317	7.443
	0.001	**0.000**	**0.000**	**0.000**	0.238	0.210
ROA	**-0.658**	**-0.832**	**-2.586**	**-3.511**	**-0.740**	**-0.946**
	0.021	**0.009**	**0.001**	**0.000**	**0.025**	**0.012**
LOSS	-0.122	0.002	0.784	0.928	-0.161	-0.010
	0.525	0.992	0.142	0.110	0.475	0.970
Nagelkerke/Adjusted R-Square	0.224	0.304	0.075	0.087	0.009	0.020

Table shows the unstandardized beta coefficients and p-values for the variables in each model using different dependent variables. P-values are presented below the beta coefficients.
(1). The main model with crash period from May to December
(2). A model with crash period from January to December
The bolded values are statistically significant (p-value < 0.05).

As a conclusion, one explanation why the predictive power is higher for all the three models when extending the crash period can be that many companies publish their annual financial report during March or early April. Thus, a lot of crashes that occur due to the financials during the next few days after the release date are not caught by the main model with shorter crash period. However, in biotechnology sector earnings announcements have much lower impact on share prices than drug development announcements (Dedman et al. 2008), referring that there is likely some other more important matters that explain this phenomenon better.

5.4.4. Pharmaceuticals compared to pure-play biotechs

The sample of this study consist of pharmaceutical companies (SIC 2834) and pure-play biotechnology companies (SIC 2833, 2835-2836). Pharmaceutical firms utilize also non-biotechnology methods in research and development, and firms can operate also in other health care areas, such as production of health care and animal products, whereas pure-play biotechnology firms focus on discovery and development of drugs (Ely et al. 2003). I test how the regression analysis results differ between these two types of biotechnology firms, and is the regression model better to explain changes in dependent variables in one type than in other type of firms.

Dividing the sample to two parts leaves 823 pharmaceutical firms and 480 pure-play biotechnology firms. With the sample of pharmaceutical firms DROP identifies at least one crash in 38.0% of the sample periods, whereas the corresponding figure is 51.3% for the pure-play biotechnology firms, indicating that the latter firms are more crash-prone than pharmaceutical companies. Further, the mean values for CRASH and NCSKEW are larger for the pure-play biotechnology firms than for the pharmaceutical firms, supporting the findings.

Table 13 presents the regression results for each crash metrics. First, the model using DROP as the dependent variable has classification accuracy of 68.2% for the pharmaceutical and 62.7% for the pure-play biotech companies (not reported), and Nagelkerke R-Square values are, respectively, 0.244 and 0.164, indicating that the accuracy and explanatory power of the model are better for pharmaceutical firms. While the main model has three statistically significant predictors, XRD, VOLATILITY, and ROA, the pharmaceutical firms have only VOLATILITY, and the pure-play biotechnology firms have only ROA as statistically significant independent variable.

Table 13 – Sensitivity analysis results: subsamples

	Beta coefficients and p-values								
	DROP			*CRASH*			*NCSKEW*		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	**1.058**	0.897	1.051	**10.356**	**10.619**	**8.296**	-0.593	-0.463	-1.045
	0.012	0.108	0.159	**0.000**	**0.000**	**0.000**	0.214	0.450	0.244
GOODWILL	-0.103	0.710	-0.946	-3.437	-1.724	-5.747	-0.862	-0.394	-1.454
	0.891	0.518	0.366	0.081	0.506	0.068	0.301	0.727	0.251
GDWLIP	0.000	-0.375	0.525	0.597	-0.317	2.775	0.250	-0.271	1.540
	0.999	0.544	0.482	0.603	0.809	0.220	0.606	0.635	0.092
XRD	**0.002**	0.001	0.002	-0.001	0.000	-0.002	-0.001	**-0.003**	0.000
	0.037	0.396	0.082	0.729	0.947	0.488	0.209	**0.030**	0.898
MTB	-0.001	0.005	-0.010	-0.004	-0.008	-0.001	0.002	-0.001	0.006
	0.771	0.452	0.185	0.759	0.661	0.977	0.678	0.911	0.467
logMCAP	-0.311	-0.173	-0.541	0.418	0.961	-0.242	**0.668**	**0.948**	0.332
	0.198	0.581	0.169	0.532	0.241	0.838	**0.018**	**0.008**	0.486
SIZE	-0.172	-0.242	-0.032	-0.524	**-0.830**	0.128	-0.222	**-0.361**	-0.015
	0.091	0.064	0.851	0.066	**0.016**	0.808	0.066	**0.016**	0.942
RCF	-0.049	-0.007	-0.087	-0.246	-0.042	-0.464	-0.105	0.022	-0.234
	0.568	0.954	0.490	0.306	0.894	0.217	0.300	0.877	0.123
VOLATILITY	**15.501**	**19.091**	12.093	**-106.3**	**-107.3**	**-101.4**	6.317	2.449	13.294
	0.001	**0.002**	0.071	**0.000**	**0.000**	**0.000**	0.238	0.733	0.105
ROA	**-0.658**	-0.467	**-1.086**	**-2.586**	**-2.486**	**-3.127**	**-0.740**	**-0.977**	-0.494
	0.021	0.207	**0.021**	**0.001**	**0.011**	**0.020**	**0.025**	**0.022**	0.361
LOSS	-0.122	-0.136	-0.099	0.784	0.657	1.061	-0.161	-0.200	-0.108
	0.525	0.587	0.752	0.142	0.315	0.260	0.475	0.482	0.776
Observations	1303	823	480	1303	823	480	1303	823	480
Nagelkerke/ Adjusted R-Square	0.224	0.244	0.164	0.075	0.077	0.064	0.009	0.014	0.004

Table shows the unstandardized beta coefficients and p-values for the variables in each model using different dependent variables and samples for pharmaceutical and biotechnology firms. P-values are presented below the beta coefficients.
(1). The main model with the whole sample
(2). A model using a subsample of pharmaceutical firms
(3). A model using a subsample of pure-play biotechnology firms
The bolded values are statistically significant (p-value < 0.05).

The second model, using CRASH as the dependent variable, has R-Square value of 0.077 for the pharmaceutical firms and 0.064 for the pure-play biotechnology firms, and F-values are, respectively, 7.8 and 4.3, indicating that the explanatory power and significance of the model are higher for pharmaceutical firms. Both regression results have VOLATILITY and ROA as statistically significant predictors similar to the main model. For the pharmaceutical firms, however, also SIZE is a statistically significant explanatory variable.

Finally, using NCSKEW as the dependent variable, Adjusted R-Square values are 0.014 for the pharmaceutical firms and 0.004 for the pure-play biotechnology firms, and F-values are, respectively, 2.2 and 1.2, indicating that the explanatory power and significance of the model are slightly higher in case of pharmaceutical firms. The results are in line with the two other models. However, both values are low, similar to the main model with NCSKEW, referring that the explanatory power and significance of the model are not great.

Overall, pharmaceutical firms seem to be less prone to stock price crashes than pure-play biotechnology firms. However, the explanatory power of the model is higher for pharmaceutical firms than for pure-play biotechnology firms, indicating that the set of independent variables in the model is better predicting crashes for pharmaceutical firms than for biotechnology firms. Further, the significance levels of the models seem to be decreasing when restricting the sample, and some changes occur in significance levels of the individual independent variables.

6. Conclusions

This thesis examines predictors for stock price crash risk in the biotechnology sector. The study is conducted as quantitative analysis applying both logistic and multiple regression to explore the predictive power of a set of variables to forecast crash risk. The crash metric of Chen et al. (2001) and the metric of Ak et al. (2016) are used in this study, as well as an additional crash measure which was created to answer some of the limitations of existing measures. The sample consists of biotechnology companies that are in the business of developing new drugs (SIC 2833-2836) covering the period from 2002 to 2016. The final sample includes 1303 observations from which 823 are pharmaceutical firms and 480 are pure-play biotechnology firms.

In relation to *the main objective* of this study, the results suggest that return on assets has a negative association with future crash risk, and volatility of a stock returns has a positive relation with future crash risk in the biotechnology sector. The results indicate that more profitable firms in relation to their total assets are less likely to experience stock price crashes, and high historical volatility of stock returns is resulting higher likelihood of future crashes. Further, to ensure the validity of the individual independent variable results, the multicollinearity is taken into consideration. The results imply that also market capitalization and size of a firms are negatively associated with crash risk, measured as a likelihood of extreme negative returns, indicating that the firms that are larger in terms of market capitalization or total assets are less prone to crashes. I suggest that smaller biotechnology firms have higher likelihood to experience stock price crashes because they have less ongoing drug development projects and hence a failure in one project is more likely to cause a stock price crash.

In relation to *the secondary objective*, the results imply that the level of consistency between the three crash risk metrics is between 70 and 80 percent, indicating that the metrics disagree on the definition of a crash or misclassify crashes in 20-30% of the cases. I find that the unstandardized crash risk metric created as part of this study and the crash risk metric of Ak et al. (2016) have the highest level of consistency, leaving the skewness measure of Chen et al. (2001) to have lower consistencies with the other two metrics. The finding is consistent with the expectation that the metrics of Chen et al. (2001) and Ak et al. (2016) have differences in firm-specific crash classifications due to the distinct definition of a crash.

As part of the study, two subsamples, pharmaceutical firms and pure-play biotechnology firms, were compared. The results suggest that pure-play biotechnology firms are more prone to stock price crashes than pharmaceutical firms. However, the explanatory power of the regression model is slightly weaker for pure-play biotechnology firms, indicating that the set of independent variables selected to the model are explaining better crashes occurring in pharmaceutical firms than in biotechnology firms.

Further, comparing the results of this study to the prior literature I find that the biotechnology industry seems to experience more crashes than other industries on average. However, to confirm the suggestion (i.e. to ensure that same methodologies are used and the results are comparable), more studies are required. Future research could compare the crash risk regression results of biotechnology sector to results of cross-industry sample in order to identify more specifically how the determinants of stock price crashes differ between the samples.

This study contributes to the literature and provide practical implications in several ways. To my knowledge, this is the first study to research determinants of crash risk in the field of biotechnology. The study provides valuable information for managers and shareholders of biotechnology firms as well as for investors who are interested in investing to biotechnology sector. A better understanding of firm characteristics that predict stock price crash risk can be helpful in option pricing, which depends on skewness and crash risk (Hutton et al. 2009). Further, as Hutton et al. (2009) suggest, the information of crash risk predictors can be utilized in portfolio planning and risk-management applications to mitigate crash risk.

The study is one of the first efforts, if not the only, to compare firm-specific classification of stock price crashes between the different crash risk metrics. Researchers can utilize the information when evaluating the validity of the results in crash risk studies. This thesis also increase the understanding of the impacts of different crash measures to the results, in respect of how they define a crash.

One of the implications of this study for research includes the new stock price crash risk metric created as part of the study. Researcher can utilize the metric as an additional crash measure to support the findings of other measures or they can use it as the main measure to identify stock price crashes. The new crash metric answers to some limitations of existing measures, and hence can match better to the demands of some researchers.

Addition to the contributions, this study has some limitations as well. The empirical evidence of this study provides information only about the associations between the predictors and the crash metrics, not causal connections. Thus, the thesis can only make suggestions about the causal connections, and further research is needed to identify the underlying phenomena of the results.

This thesis investigates how well the chosen set of variables can predict future stock price crashes. While the empirical results provide information about the association between single explanatory variables and future crash risk, the explanatory power of single predictors cannot be examined in isolation from other independent variables. Further, this study focus on biotechnology sector, and hence the results cannot be applied to other same type of industries without further research.

References

Amir, E., & Lev, B. (1996). Value-relevance of nonfinancial information: The wireless communications industry. *Journal of accounting and economics*, *22*(1), 3-30.

An, H., & Zhang, T. (2013). Stock price synchronicity, crash risk, and institutional investors. *Journal of Corporate Finance*, *21*, 1-15.

Andreou, P. C., Louca, C., & Petrou, A. P. (2017). CEO age and stock price crash risk. *Review of Finance*, *21*(3), 1287-1325.

Ak, B. K., Rossi, S., Sloan, R., & Tracy, S. (2016). Navigating stock price crashes. *The Journal of Portfolio Management*, *42*(4), 28-37.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of accounting and economics*, *24*(1), 3-37.

Callen, J. L., & Fang, X. (2015). Short interest and stock price crash risk. *Journal of Banking & Finance*, *60*, 181-194.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics*, *31*(3), 281-318.

Chen, C., Kim, J. B., & Yao, L. (2017). Earnings smoothing: Does it exacerbate or constrain stock price crash risk?. *Journal of Corporate Finance*, *42*, 36-54.

Chen, J., Hong, H., & Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of financial Economics*, *61*(3), 345-381.

Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and accrual-based earnings management in the pre-and post-Sarbanes-Oxley periods. *The accounting review*, *83*(3), 757-787.

Dedman, E., Lin, S. W. J., Prakash, A. J., & Chang, C. H. (2008). Voluntary disclosure and its impact on share prices: Evidence from the UK biotechnology sector. *Journal of Accounting and Public Policy*, *27*(3), 195-216.

DiMasi, J. A., Hansen, R. W., & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. *Journal of health economics*, *22*(2), 151-185.

DiMasi, J. A., & Grabowski, H. G. (2007). The cost of biopharmaceutical R&D: is biotech different?. *Managerial and decision Economics*, *28*(4-5), 469-479.

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: new estimates of R&D costs. *Journal of health economics*, *47*, 20-33.

Ely, K., Simko, P. J., & Thomas, L. G. (2003). The usefulness of biotechnology firms' drug development status in the evaluation of research and development costs. *Journal of Accounting, Auditing & Finance*, *18*(1), 163-196.

Francis, B., Hasan, I., & Li, L. (2016). Abnormal real operations, real earnings management, and subsequent crashes in stock prices. *Review of Quantitative Finance and Accounting*, *46*(2), 217-260.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2007). Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. *Management Science*, *53*(9), 1452-1466.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of accounting and economics*, *40*(1), 3-73.

Hamill, P. A., McIlkenny, P., & Opong, K. K. (2013). Valuation implications of pharmaceutical companies' R&D regulatory approval notifications. *The British Accounting Review*, *45*(2), 99-111.

Hand, J. R. (2001). The market valuation of biotechnology firms and biotechnology R&D. Working paper, University of North Carolina, Chapel Hill.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R^2, and crash risk. *Journal of financial Economics*, *94*(1), 67-86.

Jin, L., & Myers, S. C. (2006). R 2 around the world: New theory and new tests. *Journal of financial Economics*, *79*(2), 257-292.

Kim, J. B., Li, Y., & Zhang, L. (2011a). Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics*, *100*(3), 639-662.

Kim, J. B., Li, Y., & Zhang, L. (2011b). CFOs versus CEOs: Equity incentives and crashes. *Journal of Financial Economics*, *101*(3), 713-730.

Kim, J. B., Wang, Z., & Zhang, L. (2016). CEO overconfidence and stock price crash risk. *Contemporary Accounting Research*, *33*(4), 1720-1749.

Kim, J. B., & Zhang, L. (2016). Accounting conservatism and stock price crash risk: Firm-level evidence. *Contemporary Accounting Research*, *33*(1), 412-441.

Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance*, *43*, 1-13.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Sharma, A., & Lacey, N. (2004). Linking product development outcomes to market valuation of the firm: The case of the US pharmaceutical industry. *Journal of Product Innovation Management*, *21*(5), 297-308.

Shortridge, R. T. (2004). Market Valuation of Successful versus Non-Successful R&D Efforts in the Pharmaceutical Industry. *Journal of Business Finance & Accounting*, *31*(9-10), 1301-1325.

Watson, P. F., & Petrie, A. (2010). Method agreement analysis: a review of correct methodology. *Theriogenology*, *73*(9), 1167-1179.

Appendix A. Variable definitions

Variable	*Definition*
Dependent variables:	
DROP	Binary variable that equals 1 when a firm has experienced at least one stock price crash during the crash period from May to December.
CRASH	The negative ratio of the crash period's minimum daily return to standard deviation of previous period's daily returns.
NCSKEW	The negative skewness of daily return distribution over the crash period from May to December.
Explanatory variables:	
GOODWILL	The ratio of goodwill to total assets.
GDWLIP	Dummy variable that equals 1 if a firm has done goodwill impairments during the fiscal year, and zero otherwise. Missing values of goodwill impairments are replaced with zero.
XRD	The ratio of research and development expenses to net sales.
MTB	The ratio of market capitalization to common ordinary equity.
logMCAP	The common logarithm at market capitalization at fiscal year-end.
SIZE	The natural logarithm of total assets at fiscal year-end.
RCF	The ratio of change in cash and cash equivalent to shareholders' equity.
VOLATILITY	Standard deviation of daily returns over the previous crash period t-1.
ROA	The ratio of net income to total assets.
LOSS	Dummy variable that equals 1 if net income is negative for the fiscal year, and zero otherwise.